Shrinkage Rules for Variational Minimization Problems and Applications to Analytical Ultracentrifugation

Martin Ehler

Abstract. Finding a sparse representation of a possibly noisy signal can be modeled as a variational minimization with ℓ_q-sparsity constraints for q less than one. Especially for real-time, on-line, or iterative applications, in which problems of this type have to be solved multiple times, one needs fast algorithms to compute these minimizers.

Identifying the exact minimizers is computationally expensive. We consider minimization up to a constant factor to circumvent this limitation. We verify that q-dependent modifications of shrinkage rules provide closed formulas for such minimizers. Therefore, their computation is extremely fast. We also introduce a new shrinkage rule which is adapted to q.

To support the theoretical results, the proposed method is applied to Landweber iteration with shrinkage used at each iteration step. This approach is utilized to solve the ill-posed problem of analytic ultracentrifugation, a method to determine the size distribution of macromolecules. For relatively pure solutes, our proposed scheme leads to sparser solutions with sharper peaks, higher resolution, and smaller residuals than standard regularization for this problem.

Keywords. shrinkage, variational optimization, sparsity, frames, Fredholm integral equations.

2010 Mathematics Subject Classification. 65K10, 42C15, 65R32, 45B05.

1 Introduction

Decomposing signals into simple building blocks and reconstructing from shrinked coefficients are used in signal representation and processing, e.g., wavelet shrinkage is applied to noise and clutter reduction in speckled SAR images, cf. [38]. Statistical approaches and Bayesian objectives for noise removal make use of various shrinkage strategies, cf. [18, 24, 26, 46]. Variational models as in [9] justify

The research was funded by the Intramural Research Program of the National Institutes of Child Health and Human Development and by the Research Career Transition Awards Program 575910 of the National Institutes of Health and the German Science Foundation.

shrinkage by smoothness estimates of the unperturbed signal. Other shrinkage rules are derived from a diffusion approach in [37].

Signal approximation with sparsity constraints leads to variational minimization problems, and the denoising approach in [9] is a particular case. The expression to be minimized is a sum of an approximation error and a penalty term which involves weighted ℓ_q-constraints, see Section 2.1. In [14], iteratively shrinking coefficients of an orthonormal basis expansion provides a sequence converging towards the minimizer. The method covers the convex case $q \in [1,2]$, but sparse signal representation, coding, signal analysis, and the treatment of operator equations require the consideration of redundant basis-like systems and the nonconvex case $q \in [0,1)$ as well, see for instance [2, 10, 12, 13, 23, 28, 42] and references therein. By using hard-shrinkage, the algorithm in [14] converges towards a local minimum for $q = 0$, cf. [4]. Under the restricted isometry property (RIP) [8, 17], the iteration converges towards the exact minimum, cf. [3]. However, RIP does not hold in many situations and therefore the local minimum could be far off the global minimum. The approach does not cover $q \in (0,1)$, and, for applications where computation time is crucial, a faster algorithm is desirable.

In the present paper, we obtain complementary results for $q \in [0,1)$ in terms of minimization up to a constant factor. In fact, we verify that such a minimization can be derived from q-dependent modifications of shrinkage rules. This means we have a closed formula for these minimizers, which allows for a fast computation. We also introduce new shrinkage rules which are adapted to q. We then propose a Landweber iteration with these new shrinkage strategies applied in each step to treat sparsity constraints for $q \in (0,1)$, cf. [4, 14] for soft- and hard-shrinkage. This approach is then applied to the ill-posed problem of sedimentation velocity analytical ultracentrifugation, a method to determine the size distribution of interacting macromolecules [11, 41]. Its physical model leads to a Fredholm integral equation that needs to be regularized. For highly pure monomers, the solution is expected to be highly sparse with few sharp peaks. Our numerical experiments suggest that our proposed iterative shrinkage scheme leads to sharper peaks, fewer nonzero entries, and smaller residuals. Thus, it provides a useful add-on to standard analytical ultracentrifugation analysis, cf. [7, 41].

The outline is as follows: In Section 2, we present the variational problems under consideration and we recall the concept of frames. We introduce shrinkage rules in Section 3. The main results about minimization up to a constant factor are presented in Section 4, and in Section 5 we apply the results to sparse signal representation. We introduce a new family of shrinkage rules in Section 6. The modified Landweber iteration is explicitly introduced in Section 7, where we also present numerical results about sedimentation velocity analytical ultracentrifugation. Conclusions are given in Section 8.

2 Variational Problems and Frames

2.1 Variational Minimization Problems

Let L be a bounded operator between two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}'$, and let $\{\tilde{f}_n\}_{n\in\mathcal{N}}$ be a countable collection in $\mathcal{H}$. Given $h \in \mathcal{H}'$, we consider the minimization problem

$$\min_{g\in\mathcal{H}} \big(\|h - Lg\|^2_{\mathcal{H}'} + \sum_{n\in\mathcal{N}} \alpha_n |\langle g, \tilde{f}_n\rangle|^q\big), \tag{2.1}$$

where $q \in (0, 2]$, $(\alpha_n)_{n\in\mathcal{N}}$ is a sequence of nonnegative numbers, and $\langle\cdot,\cdot\rangle$ denotes the inner product on $\mathcal{H}$. This makes also sense for $q = 0$ with the convention $0^0 = 0$, and the penalty term then counts the nonzero entries of $(\langle g, \tilde{f}_n\rangle)_{n\in\mathcal{N}}$ weighted by $(\alpha_n)_{n\in\mathcal{N}}$. For $\mathcal{H} = \mathcal{H}'$ and $L = \mathrm{id}_{\mathcal{H}}$, problem (2.1) is relevant in wavelet based signal denoising. There, $\{\tilde{f}_n\}_{n\in\mathcal{N}}$ is a wavelet system, and the sparsity constraint on the right hand side of (2.1) is related to the Besov regularity of the signal to be recovered, see [9] for details. Our approach is neither restricted to L being the identity nor must L be injective. However, we assume that it has a bounded pseudo inverse, i.e. there is a bounded operator $L^{\#} : \mathcal{H}' \mapsto \mathcal{H}$ such that $LL^{\#}L = L$. Thus, we first consider well-posed problems and are therefore more restrictive than in [14]. Nevertheless, we address ill-posed problems in Section 7 by extending the iterative shrinkage procedure introduced in [14].

The sequence $(\alpha_n)_{n\in\mathcal{N}}$ is a collection of variable parameters which must be fitted to h and L. If all components of $\alpha_n = \alpha$ are identical, then

$$\alpha \mapsto \big(\|h - Lg^{\alpha}\|^2_{\mathcal{H}'}, \sum_{n\in\mathcal{N}} |\langle g^{\alpha}, \tilde{f}_n\rangle|^q\big) \tag{2.2}$$

is considered as a curve in $\mathbb{R}^2$, where g^{α} is a minimizer of (2.1), and one finally chooses α according to a point of maximal curvature, see [33] and [36] for the *L-curve* and *H-curve criterion*, respectively. It requires to compute minimizers g^{α} for many different values of α, and g^{α} must be efficiently computable. This is another motivation for avoiding costly iterative minimization schemes beside real-time and on-line applications.

Handling nonstationary noise requires that $(\alpha_n)_{n\in\mathcal{N}}$ depends on n, but it is often still reasonable to assume that there are positive constants a and b such that

$$a \le \alpha_n \le b, \text{ for all } n \in \mathcal{N}. \tag{2.3}$$

2.2 Bi-frames

The singular value decomposition of L is considered in [35] to address $q \in [0, 1)$. The system $\{\tilde{f}_n\}$ in (2.1) is supposed to be an orthonormal basis for $\mathcal{H}$ which

diagonalizes L. However, diagonalizing L can be extremely difficult in practical applications. We will consider redundant basis-like systems, and L is not required to be diagonalized: a countable collection $\{f_n\}_{n\in\mathcal{N}}$ in $\mathcal{H}$ is a *frame* if there are two positive constants A, B such that

$$A\|g\|_{\mathcal{H}}^2 \leq \sum_{n\in\mathcal{N}} |\langle g, f_n\rangle|^2 \leq B\|g\|_{\mathcal{H}}^2, \text{ for all } g \in \mathcal{H}. \tag{2.4}$$

If $\{f_n\}_{n\in\mathcal{N}}$ is a frame, then its *synthesis operator*

$$F : \ell_2(\mathcal{N}) \to \mathcal{H}, \quad (c_n)_{n\in\mathcal{N}} \mapsto \sum_{n\in\mathcal{N}} c_n f_n, \tag{2.5}$$

is onto. Each $g \in \mathcal{H}$ then has a series expansion, but we still have to find its coefficients. The synthesis operator's adjoint

$$F^* : \mathcal{H} \to \ell_2(\mathcal{N}), \quad g \mapsto (\langle g, f_n\rangle)_{n\in\mathcal{N}} \tag{2.6}$$

is called *analysis operator*, $S = FF^*$ is invertible, and $\{S^{-1}f_n\}_{n\in\mathcal{N}}$ is called *canonical dual frame* and expands

$$g = \sum_{n\in\mathcal{N}} \langle g, S^{-1}f_n\rangle f_n, \text{ for all } g \in \mathcal{H}.$$

The inversion of S can be difficult, and, since F need not be injective, there could be 'better' coefficients than $\langle g, S^{-1}f_n\rangle$. This motivates the following: two frames $\{f_n\}_{n\in\mathcal{N}}$ and $\{\tilde{f}_n\}_{n\in\mathcal{N}}$ are called a *pair of dual frames* (or a *bi-frame*) if

$$g = \sum_{n\in\mathcal{N}} \langle g, \tilde{f}_n\rangle f_n, \text{ for all } g \in \mathcal{H}, \tag{2.7}$$

i.e., $F\widetilde{F}^* = \mathrm{id}_{\mathcal{H}}$, where $\widetilde{F}^*$ is the dual frame's analysis operator. For instance, the canonical dual of a wavelet frame may not have the wavelet structure as well, but it can possibly replaced by an alternative dual wavelet frame, cf. [16,19,21,22,31,40] and references therein.

Throughout the paper, we suppose that $\{f_n\}_{n\in\mathcal{N}}$ and $\{\tilde{f}_n\}_{n\in\mathcal{N}}$ are a **bi-frame for** $\mathcal{H}$.

3 Shrinkage Rules

To solve (2.1), shrinkage plays a crucial role. Following ideas in [43], we call a function $\varrho : \mathbb{C}\times\mathbb{R}_{\geq 0} \to \mathbb{C}$ a *shrinkage rule* if there are constants $C_1, C_2, \rho, D > 0$

such that both conditions

$$|x - \varrho(x,\alpha)| \le C_1 \min(|x|,\alpha), \text{ for all } \alpha \ge 0,\, x \in \mathbb{C}, \tag{3.1}$$

$$|\varrho(x,\alpha)| \le C_2 |x| \big|\frac{x}{\alpha}\big|^{\rho}, \text{ for all } \alpha > 0,\, |x| \le D\alpha, \tag{3.2}$$

are satisfied. While (3.1) forces $\varrho(x,\alpha)$ to be close to x for small α, condition (3.2) means that $\varrho(x,\alpha)$ has sufficient decay as x goes to zero. A shrinkage rule ϱ is called a *thresholding rule* if there is a constant $C_3 > 0$ such that $|x| \le C_3\alpha$ implies $\varrho(x,\alpha) = 0$. A thresholding rule allows for $\rho = \infty$ in (3.2), where we use $a^\infty = 0$ if $0 \le a < 1$. We will recall a few common shrinkage rules and we restrict us to $x \in \mathbb{R}$, see also Figure 1: *Soft-shrinkage* is given by $\varrho_s(x,\alpha) = (x - \text{sign}(x)\alpha)\mathbf{1}_{\{|x|>\alpha\}}$. Contrary to soft- and *hard-shrinkage* $\varrho_h(x,\alpha) = x\mathbf{1}_{\{|x|>\alpha\}}$, the *nonnegative garotte-shrinkage* rule $\varrho_g(x,\alpha) = (x - \frac{\alpha^2}{x})\mathbf{1}_{\{|x|>\alpha\}}$ is continuous and large coefficients are not changed much. It has been successfully applied to image denoising in [29]. Similar properties has *hyperbolic-shrinkage* $\varrho^{hy}(x,\alpha) = \text{sign}(x)\sqrt{x^2-\alpha^2}\mathbf{1}_{\{|x|>\alpha\}}(x)$, cf. [43].

The *n-degree garotte-shrinkage* rule is given by $\varrho^n(x,\alpha) = \frac{x^{2n+1}}{x^{2n}+\alpha^{2n}}$, see [43]. For $k \in \mathbb{N}$, the twice differentiable rule

$$\varrho_k(x,\alpha) = \begin{cases} \frac{x^{2k+1}}{(2k+1)\alpha^{2k}}, & |x| \le \alpha \\ x - \text{sign}(x)(\alpha - \frac{\alpha}{2k+1}), & |x| > \alpha \end{cases} \tag{3.3}$$

is considered in [46]. Both rules are shrinkage rules with $\rho = 2k = 2n$ and $C_2 = 1$. The rules $\varrho(x,\alpha) = x\big(1 - \sqrt{\frac{\alpha^2}{\alpha^2+2x^2}}\big)$ and $\varrho(x,\alpha) \approx x\exp(-0.2\frac{\alpha^8}{x^8})$ are based on diffusion, see [37]. One verifies that both are shrinkage rules with $\rho = 1$, and we refer to them as diffusion 1 and 2 in Figure 1.

Bruce and Gao proposed *firm-shrinkage*

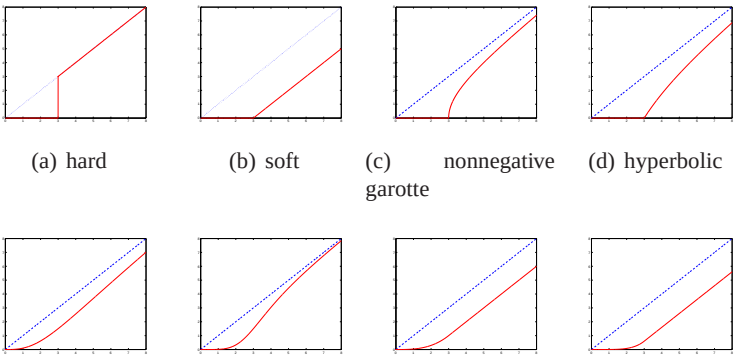
$$\varrho_f(x,\alpha_1,\alpha_2) = x\mathbf{1}_{\{|x|>\alpha_2\}} + \text{sign}(x)\frac{\alpha_2(|x|-\alpha_1)}{\alpha_2-\alpha_1}\mathbf{1}_{\{\alpha_1\le|x|\le\alpha_2\}}$$

in [30]. For fixed α_1, the mapping $(x,\alpha) \mapsto \varrho_f(x,\alpha_1,\alpha)$ is a thresholding rule.

4 Main Results

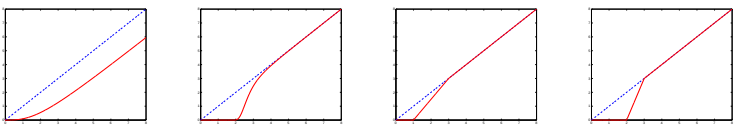
For $q \in [0,2]$, let $\ell_q^{(\alpha_n)}(\mathcal{N})$ denote the weighted $\ell_q(\mathcal{N})$-space, i.e., the space of complex-valued sequences $(\omega_n)_{n\in\mathcal{N}}$ such that $\|\omega\|_{\ell_q^{(\alpha_n)}}^q := \sum_{n\in\mathcal{N}} \alpha_n|\omega_n|^q$ is finite. One observes that $\sum_{n\in\mathcal{N}} \alpha_n|\langle g, \tilde{f}_n\rangle|^q = \|\widetilde{F}^* g\|_{\ell_q^{(\alpha_n)}}^q$, and to shorten notation, we denote

$$\mathcal{J}_q(h,g) = \|h - Lg\|_{\mathcal{H}'}^2 + \|\widetilde{F}^* g\|_{\ell_q^{(\alpha_n)}}^q.$$



(a) hard (b) soft (c) nonnegative garotte (d) hyperbolic

(e) 1-degree garotte (f) 2-degree garotte (g) $k = 1$-shrinkage (h) $k = 2$-shrinkage



(i) diffusion 1 (j) diffusion 2 (k) firm, $\alpha_1 = 1$ (l) firm, $\alpha_1 = 2$

Figure 1. Shrinkage rules $\varrho(x, \alpha)$, for $\alpha = 3$. (a) is not continuous. (b)-(d),(k), and (l) are continuous but not differentiable. (e)-(j) are smooth

The idea for the following main result is to replace a shrinkage rule $\varrho(x, \alpha)$ by its q-dependent expression $\varrho(x, \alpha|x|^{q-1})$. Due to (3.1), it vanishes as $x \neq 0$ goes to 0, and we apply $\varrho(x, \alpha|x|^{q-1}) = 0$ for $x = 0$. If $\rho = \infty$, we use $\frac{1}{\rho} = 0$. Since $\varrho_g(x, \alpha) = \varrho_s(x, \alpha^2|x|^{-1})$, the nonnegative garotte is q-dependent soft-shrinkage for $q = 0$ and α replaced by α^2. It turns out that q-dependent shrinkage expressions provide minimizers of (2.1) up to a constant factor:

Theorem 4.1. *Let ϱ be a shrinkage rule with $\rho \in [\frac{1}{2}, \infty]$. Suppose that $\widetilde{F}^* L^\# LF$ is bounded on $\ell^{(\alpha_n)}_{1/\rho}(\mathcal{N})$. Let $q = \frac{1}{\rho}$, then there is a constant $C > 0$ such that for all $h \in \mathrm{range}(L)$, and for all $g \in \mathcal{H}$*

$$\mathcal{J}_q(h, \hat{g}) \leq C \mathcal{J}_q(h, g),$$

where $\hat{g} = L^\# LF\varrho(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$ with $v = \widetilde{F}^* L^\# h$.

If $\widetilde{F}^* F$ is also bounded on $\ell^{(\alpha_n)}_{1/\rho}(\mathcal{N})$, one can choose $\hat{g} = F\varrho(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$. If (2.3) *holds, then the statements extend to all $q \in [\frac{1}{\rho}, 2]$, and C is independent of q.*

Remark 4.2. If the bi-frame is biorthogonal and $F\ell_{1/\rho}^{(\alpha_n)} \subset \text{range}(L^\# L)$, then $\widetilde{F}^* L^\# L F = \text{id}_{\ell_{1/\rho}^{(\alpha_n)}}$, because $\widetilde{F}^* F$ is the identity and $L^\# L$ is the identity on its range. The boundedness condition is then trivially satisfied as it is for finite $\mathcal{N}$.

To prove Theorem 4.1, we consider a decoupled minimization problem: given $v \in \ell_2(\mathcal{N})$, we try to minimize

$$\mathcal{I}_q(v,\omega) = \big(\|v-\omega\|_{\ell_2}^2 + \sum_{n\in\mathcal{N}} \alpha_n |\omega_n|^q\big) \tag{4.1}$$

over $\omega \in \ell_2(\mathcal{N})$. It turns out that minimizing (2.1) and (4.1) up to a constant factor are equivalent:

Proposition 4.3. *Given $q \in [0,2]$, suppose that $\widetilde{F}^* L^\# L F$ is bounded on $\ell_q^{(\alpha_n)}(\mathcal{N})$. For $h \in \text{range}(L)$, let $v = \widetilde{F}^* L^\# h$. If $\hat{\omega}$ minimizes* (4.1) *up to a constant factor, then $\hat{g} = L^\# L F \hat{\omega}$ minimizes* (2.1) *up to a constant factor.*

If $\widetilde{F}^* F$ is bounded on $\ell_q^{(\alpha_n)}(\mathcal{N})$, one may also choose $\hat{g} = F\hat{\omega}$. The reverse implication holds for $\hat{\omega} = \widetilde{F}^* L^\# L \hat{g}$ and $\hat{\omega} = \widetilde{F}^* \hat{g}$, respectively.

Given a parameter set Γ and two expressions $(a_\tau)_{\tau\in\Gamma}$ and $(b_\tau)_{\tau\in\Gamma}$ such that there is a constant $C > 0$ with $a_\tau \le C b_\tau$ for all $\tau \in \Gamma$, we write $a_\tau \lesssim b_\tau$ in the following proof.

Proof of Proposition 4.3. Let $\hat{\omega}$ minimize (4.1) up to a constant factor, i.e., $\mathcal{I}_q(v,\hat{\omega}) \lesssim \mathcal{I}_q(v,\omega)$, for all $\omega \in \ell_2(\mathcal{N})$. Since $F\widetilde{F}^* = \text{id}_{\mathcal{H}}$ and since $LL^\# L = L$ yields $LL^\# h = h$, we have $h = LFv$. Applying $LL^\# L = L$ implies $L\hat{g} = LF\hat{\omega}$, which leads to

$$\mathcal{J}_q(h,\hat{g}) = \|LFv - LF\hat{\omega}\|_{\mathcal{H}'}^2 + \|\widetilde{F}^* L^\# LF\hat{\omega}\|_{\ell_q^{(\alpha_n)}}^q.$$

Since $LF : \ell_2 \mapsto \mathcal{H}'$ is bounded and due to the boundedness of $\widetilde{F}^* L^\# LF$ on $\ell_q^{(\alpha_n)}$, this implies $\mathcal{J}_q(h,\hat{g}) \lesssim \mathcal{I}_q(v,\hat{\omega})$. Since $\hat{\omega}$ minimizes (4.1) up to a constant factor, we have $\mathcal{J}_q(h,\hat{g}) \lesssim \mathcal{I}_q(v, \widetilde{F}^* L^\# L g)$, for all $g \in \mathcal{H}$. By applying that $\widetilde{F}^* L^\#$ is bounded and that $F\widetilde{F}^* = \text{id}_{\mathcal{H}}$, we obtain, for all $g \in \mathcal{H}$,

$$\begin{aligned}\mathcal{J}_q(h,\hat{g}) &\lesssim \|\widetilde{F}^* L^\# h - \widetilde{F}^* L^\# L g\|_{\ell_2}^2 + \|\widetilde{F}^* L^\# LF\widetilde{F}^* g\|_{\ell_q^{(\alpha_n)}}^q \\ &\lesssim \|h - Lg\|_{\ell_2}^2 + \|\widetilde{F}^* L^\# LF\widetilde{F}^* g\|_{\ell_q^{(\alpha_n)}}^q \lesssim \mathcal{J}_q(h,g),\end{aligned}$$

where we have used that $\widetilde{F}^* L^\# LF$ is bounded on $\ell_q^{(\alpha_n)}$.

Analogous arguments can be applied to the case $\hat{g} = F\hat{\omega}$, and the reverse implications follow in a similar way. □

Next, we obtain a solution of the discrete problem (4.1).

Proposition 4.4. *Let ϱ be a shrinkage rule with $\rho \in [\frac{1}{2}, \infty]$. Then there is a constant $C > 0$ such that for all $q \in [\frac{1}{\rho}, 2]$, for all $v \in \ell_2(\mathcal{N})$, and for all $\omega \in \ell_2(\mathcal{N})$,*

$$\mathcal{I}_q(v, \hat{\omega}) \leq C\mathcal{I}_q(v, \omega),$$

where $\hat{\omega} = \varrho(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$.

Remark 4.5. The exact minimizer of (4.1) for $q = 2$ is known to be $\big(\frac{1}{1+\alpha_n} v_n\big)_{n\in\mathcal{N}}$. However, $(x, \alpha) \mapsto \frac{1}{1+\alpha} x$ is not a shrinkage rule since (3.2) is violated. On the other hand, the rule $\varrho(x, \alpha) = \frac{1}{1+\frac{\alpha}{|x|}} x$ is a shrinkage rule with constant $\rho = 1$. The q-dependent expression $\varrho(x, \alpha|x|^{q-1})$ for $q = 2$ then yields the exact minimizer. In this sense the exact minimizer for $q = 2$ is still derived from shrinkage.

Proof of Proposition 4.4. First, we consider $\frac{1}{2} \leq \rho < \infty$. Due to (3.1), the sequence $\varrho(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$ is indeed contained in $\ell_2(\mathcal{N})$. Adapting results in [9] to our setting yields that the hard-shrinked sequence $\varrho_h(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$ minimizes (4.1) up to a constant factor. By using the short-hand notation

$$\begin{aligned} K_n &:= \big|v_n - \varrho_h(v_n, \alpha_n|v_n|^{q-1})\big|^2 + \alpha_n\big|\varrho_h(v_n, \alpha_n|v_n|^{q-1})\big|^q, \\ G_n &:= |v_n - \varrho(v_n, \alpha_n|v_n|^{q-1})|^2 + \alpha_n|\varrho(v_n, \alpha_n|v_n|^{q-1})|^q, \end{aligned}$$

we consider each n in the sequence norms separately. We aim to verify $G_n \lesssim K_n$ independently of n. For $v_n = 0$, we have $G_n = K_n$. Now, we suppose $v_n \neq 0$. Since (3.2) gets weaker as ρ and D decrease, we may assume that $q = \frac{1}{\rho}$ and $D \leq 1$. Case 1: For $|v_n| \leq D\alpha_n|v_n|^{q-1}$, (3.1) and (3.2) with $\rho = \frac{1}{q}$ yield

$$\begin{aligned} G_n &\leq C_1^2|v_n|^2 + \alpha_n C_2^q |v_n|^q \frac{|v_n|}{\alpha_n|v_n|^{q-1}} \\ &\leq C_1^2|v_n|^2 + C_2^q|v_n|^2 \lesssim |v_n|^2 = K_n. \end{aligned}$$

Case 2: For $|v_n| > D\alpha_n|v_n|^{q-1}$, we have $1/D > \alpha_n|v_n|^{q-2}$, and the estimate (3.1) yields

$$\begin{aligned} G_n &\leq C_1^2(\alpha_n|v_n|^{q-1})^2 + \alpha_n\big(|v_n| + C_1 \min(|v_n|, \alpha_n|v_n|^{q-1})\big)^q \\ &\leq C_1^2\alpha_n|v_n|^q\alpha_n|v_n|^{q-2} + \alpha_n|v_n|^q(1 + C_1\alpha_n|v_n|^{q-2})^q \\ &\leq C_1^2\alpha|v_n|^q\tfrac{1}{D} + (1 + C_1/D)^q\alpha_n|v_n|^q \lesssim \alpha_n|v_n|^q \leq K_n/D. \end{aligned}$$

Hence, $G_n \lesssim K_n$ holds in both cases. Similar arguments verify the statement for $\rho = \infty$. □

Our main result follows from combining both propositions:

Proof of Theorem 4.1. According to Proposition 4.4, $\varrho(v_n, \alpha_n |v_n|^{q-1})_{n\in\mathcal{N}}$ is a minimizer of (4.1) up to a constant factor, where $v = \widetilde{F}^* L^\# f$. For $q = \frac{1}{\rho}$, Proposition 4.3 then implies Theorem 4.1. If (2.3) holds, $\widetilde{F}^* L^\# LF$ and $\widetilde{F}^* F$ are bounded on $\ell_2^{(\alpha_n)}$. Interpolation between $\ell_{1/\rho}^{(\alpha_n)}$ and $\ell_2^{(\alpha_n)}$ yields uniform boundedness on $\ell_q^{(\alpha_n)}$, for $q \in [\frac{1}{\rho}, 2]$. □

Remark 4.6. We did not use the Hilbert space structure of $\mathcal{H}'$ and in fact Theorem 4.1 still holds if $\mathcal{H}'$ is a (quasi) Banach space.

5 Sparse Approximation

Given $h \in \mathcal{H}$ (possibly noisy) and a frame $\{f_n\}_{n\in\mathcal{N}}$ for $\mathcal{H}$, an important problem in sparse signal representation is to find the minimizer of

$$\min_{\omega\in\ell_2} \|\omega\|_{\ell_q} \text{ subject to } F\omega \approx h, \tag{5.1}$$

for $q \in [0, 1)$. Under additional requirements on $\{f_n\}_{n\in\mathcal{N}}$ and h, the solution for $q \in [0, 1)$ can be obtained from solving the much simpler convex problem with $q = 1$, cf. [8,17]. However, these results are limited to finite $\mathcal{N}$, and the additional requirements are not satisfied in many situations.

The problem (5.1) is often replaced by a variational formulation, and one seeks to minimize

$$\mathcal{K}_q(h, \omega) = \|h - F\omega\|_{\mathcal{H}}^2 + \sum_{n\in\mathcal{N}} \alpha_n |\omega_n|^q$$

over $\omega \in \ell_2(\mathcal{N})$. For finite $\mathcal{N}$, ℓ_q-basis-pursuit as in [44], for instance, solves (5.1) by minimizing $\|F^\# h\|_{\ell_q}$ over all pseudo inverses $F^\#$. The associated variational formulation is

$$\min_{F^\#} \big(\min_{\omega\in\ell_2} \big(\|F^\# h - \omega\|_{\ell_2} + \sum_{n\in\mathcal{N}} \alpha_n |\omega_n|^q \big) \big).$$

We do not require $\mathcal{N}$ to be finite, and instead of minimizing over $F^\#$, we suppose to have a particular pseudo inverse $\widetilde{F}^*$ being the analysis operator of a dual frame $\{\tilde{f}_n\}_{n\in\mathcal{N}}$ such that $\widetilde{F}^* F$ is bounded on $\ell_q^{(\alpha_n)}(\mathcal{N})$:

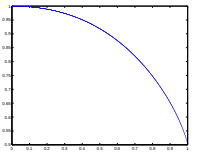


Figure 2. The curve c_q, for $q \in [0,1]$, is continuous and continuation yields $c_0 = 1$ and $c_1 = \frac{1}{2}$.

Theorem 5.1. *Given a bi-frame $\{f_n\}_{n\in\mathcal{N}}$ and $\{\tilde{f}_n\}_{n\in\mathcal{N}}$, let ϱ be a shrinkage rule with $\rho \in [\frac{1}{2}, \infty]$. Suppose that $\widetilde{F}^*F$ is bounded on $\ell^{(\alpha_n)}_{1/\rho}(\mathcal{N})$. Let $q = \frac{1}{\rho}$, then there is a constant $C > 0$ such that for all $h \in \mathcal{H}$ and for all $\omega \in \ell_2(\mathcal{N})$*

$$\mathcal{K}_q(h, \hat{\omega}) \leq C\mathcal{K}_q(h, \omega),$$

where $\hat{\omega} = \widetilde{F}^*F\varrho(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$ with $v = \widetilde{F}^*h$ or $\hat{\omega} = \varrho(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$. If (2.3) *holds, then the statement extends to all $q \in [\frac{1}{\rho}, 2]$, and C is independent of q.*

Remark 5.2. For sufficiently smooth wavelet bi-frames with vanishing moments, the operator $\widetilde{F}^*F$ is bounded on $\ell^{(\alpha_n)}_{1/\rho}$ provided that $(\alpha_n)_{n\in\mathcal{N}}$ satisfies (2.3), cf. [20].

Proof. We replace $\mathcal{H}$, $\mathcal{H}'$, L, $L^\#$, and the bi-frame $\{f_n\}_{n\in\mathcal{N}}$, $\{\tilde{f}_n\}_{n\in\mathcal{N}}$ in (2.1) by $\ell_2(\mathcal{N})$, $\mathcal{H}$, F, $\widetilde{F}^*$, and the canonical basis $\{e_n\}_{n\in\mathcal{N}}$ for $\ell_2(\mathcal{N})$, respectively. The condition on $\widetilde{F}^*L^\#LF$ in Theorem 4.1 becomes '$\widetilde{F}^*F$ is bounded on $\ell^{(\alpha_n)}_{1/\rho}(\mathcal{N})$', and Theorem 4.1 implies Theorem 5.1. □

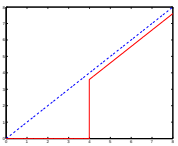
6 Explicit Shrinkage Rules Between Hard- and Soft-Shrinkage

This section is dedicated to finding a family of shrinkage rules which is adapted to q in (4.1). For $q \in [0,1)$, we will use $c_q = 2^{q-2}\frac{(2-q)^{2-q}}{(1-q)^{1-q}}$. It is monotonically decreasing with $c_0 = 1$, and continuous extension yields $c_1 = \frac{1}{2}$, see Figure 2.

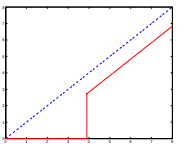
Due to [1], the exact minimizer of (4.1) is sandwiched between soft- and hard-shrinkage. Let us introduce the new shrinkage rule

$$\varrho^{(q)}_{h,s}(x, \alpha) = (x - \text{sign}(x)qc_q\alpha)\mathbf{1}_{\{|x|>\alpha c_q\}}, \tag{6.1}$$

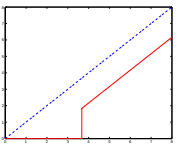
see Figure 3.



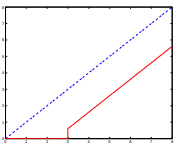
(a) $q = 0.1$



(b) $q = 0.3$

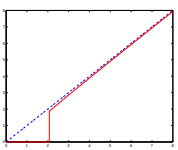


(c) $q = 1/2$

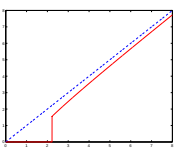


(d) $q = 0.8$

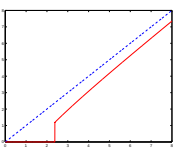
Figure 3. Shrinkage rule $\varrho_{h,s}^{(q)}(x, \alpha)$ for $\alpha = 4$. It tends to hard-shrinkage for $q \searrow 0$. Soft-shrinkage is approximated by $q \nearrow 1$.



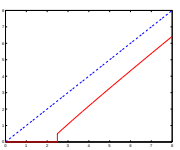
(a) $q = 0.1$



(b) $q = 0.3$



(c) $q = 1/2$



(d) $q = 0.8$

Figure 4. q-dependent expression $\varrho_{h,s}^{(q)}(x, \alpha|x|^{q-1})$ for $\alpha = 4$. According to (6.2), the smaller q the faster tends this expression to x.

One easily verifies that $\varrho_{h,s}^{(q)}(x, \alpha)$ has a jump of size $(1 - q)c_q\alpha$ and, for $|x| > c_q\alpha$, we have

$$|x - \varrho_{h,s}^{(q)}(x, \alpha)| = qc_q\alpha.$$

On the other hand, the q-dependent expression $\varrho_{h,s}^{(q)}(x, \alpha|x|^{q-1})$, see Figure 4, has a jump of size $(1 - q)(c_q\alpha)^{\frac{1}{2-q}}$ and, for $|x| > (c_q\alpha)^{\frac{1}{2-q}}$, we obtain

$$|x - \varrho_{h,s}^{(q)}(x, \alpha|x|^{q-1})| = qc_q\alpha|x|^{q-1}. \tag{6.2}$$

Hence, for $q \in [0, 1)$, the difference goes to zero as x goes to infinity.

The following theorem says that the new rule (6.1) is well adapted to $q \in [0, 1]$:

Theorem 6.1. *The sequence $\varrho_{h,s}^{(q)}(v_n, \alpha_n|v_n|^{q-1})_{n\in\mathcal{N}}$ is an exact minimizer of* (4.1) *at the endpoints $q = 0$, $q = 1$. It minimizes* (4.1) *up to a constant factor in between, and it coincides with the exact minimizer on $\{n \in \mathcal{N} : |v_n| < c_q^{\frac{1}{2-q}}\alpha_n^{\frac{1}{2-q}}\}$.*

Proof. Soft-shrinkage $\varrho_s(v_n, \frac{\alpha_n}{2})_{n\in\mathcal{N}}$ is the exact minimizer of (4.1), for $q = 1$, cf. [9]. Note that

$$\varrho_s(v_n, \frac{\alpha_n}{2}) = \varrho_{h,s}^{(1)}(v_n, \alpha_n), \quad \text{for all } n \in \mathcal{N}.$$

The exact minimizer for $q = 0$ is hard-shrinkage $\varrho_h(v_n, \sqrt{\alpha_n})_{n\in\mathcal{N}}$, see [35], and we have the identity

$$\varrho_h(v_n, \sqrt{\alpha_n}) = \varrho_{h,s}^{(0)}(v_n, \alpha_n|v_n|^{-1}) \quad \text{for all } n \in \mathcal{N}.$$

The shrinkage rule $\varrho_{h,s}^{(q)}$ satisfies (3.2) for $\rho = \infty$. Hence due to Proposition 4.4, it minimizes (4.1) up to a constant factor.

We have $\varrho_{h,s}^{(q)}(v_n, \alpha_n|v_n|^{q-1}) = 0$ iff $|v_n| \leq c_q\alpha_n|v_n|^{q-1}$. Since $|v_n| \leq c_q\alpha_n|v_n|^{q-1}$ is equivalent to $|v_n|^{2-q} \leq c_q\alpha_n$, it is also equivalent to $|v_n| \leq c_q^{\frac{1}{2-q}}\alpha_n^{\frac{1}{2-q}}$, for $q \in (0, 1)$. According to the results in [35], see also [1], each exact minimizer $(\hat{\omega}_n)_{n\in\mathcal{N}}$ satisfies $\hat{\omega}_n = 0$ for $|v_n| < c_q^{\frac{1}{2-q}}\alpha_n^{\frac{1}{2-q}}$. □

Due to Theorem 6.1, the rule $\varrho_{h,s}^{(q)}$ is an adaptation to $q \in [0, 1]$. This might also be useful for parameter fitting: While $\alpha = (\alpha_n)_{n\in\mathcal{N}}$ can be fitted to f and L by considering (2.2), the new family $\varrho_{h,s}^{(q)}$ provides additional flexibility to optimize the choice of q as well. One optimizes $\alpha = \alpha(q)$ as in (2.2), one may then vary $q \in [0, 1]$ and may optimize this sparsity parameter by analyzing the univariate curve $\alpha(q)$.

7 Iterative Shrinkage Strategies

In the present section, the derived shrinkage strategies in Section 6 are applied to inverse problems. We slightly change our perspective and consider the problem

$$\arg\min_g \big(\|f - \mathcal{T}g\|\big) \tag{7.1}$$

in which the operator $\mathcal{T}$ does not have a bounded pseudo inverse or the norm is extremely big. Such an ill-posed problem needs regularization. During the last decade, regularization with sparsity constraints has attracted significant attention, see, for instance, [5, 6, 14, 15, 25, 27, 35, 39]. One solves

$$\arg\min_g \big(\|f - \mathcal{T}g\|^2 + \alpha\phi(g)\big), \tag{7.2}$$

where ϕ is a measure of the sparsity of g in some chosen dictionary, and the non-negative regularization parameter α weights the sparsity term.

7.1 Landweber Iteration with Shrinkage

A shrinked Landweber iteration has been developed in [14] to minimize (7.2) for the term $\phi(g) = \sum_k |g_k|^q$, provided that $1 \leq q \leq 2$ and $(g_k)_k$ are the coefficients

for g's representation in an orthonormal basis. To reduce notation, let us assume that f and g are already discretized and hence are just sequences. For $q = 1$, the term $\phi(g)$ enforces sparsity. The minimization (7.2) then is well-posed, and the iteration given by

$$g^0 = 0, \tag{7.3}$$

$$g^{j+1} = S_\alpha(g^j + \mathcal{T}^* f - \mathcal{T}^* \mathcal{T} g^j), \quad \text{where } S_\alpha(x)_n = \varrho_s(x_n, \alpha), \tag{7.4}$$

converges towards the minimizer of (7.2), see [14]. Soft-shrinkage occurs in this iterative scheme, because it is the exact minimizer of (4.1). To address other $1 < q \leq 2$ in (7.2) with $\phi(g) = \sum_k |g_k|^q$, we need to apply the shrinkage rule that corresponds to the exact minimizer of (4.1) for this particular q. On the other hand, it is shown in [45] that (7.2) with the stronger sparsity requirements $0 < p < 1$ is still well-posed and a regularization of the original problem (7.1). However, the nonconvexity of $\phi(g)$ in this case makes it difficult to design a numerically attractive algorithm for the actual minimization. The q-dependent expression of the shrinkage rule $\varrho^{(q)}$ is not the exact minimizer of (4.1), but still a minimizer up to a constant factor. Motivated by the results in Section 6, we propose to replace soft-shrinkage with the q-dependent expression of $\varrho^{(q)}$, i.e., to replace S_α in (7.4) with

$$\tilde{S}_\alpha(x)_n = \varrho^{(q)}(x_n, \alpha |x_n|^{q-1}).$$

This modified scheme is known to converge for $q = 0$ and $q = 1$, cf. [4, 14]. It is thus reasonable to believe that it also converges for $0 < q < 1$, which is supported by numerical experiments.

7.2 Analytic Ultracentrifugation

Sedimentation velocity analytical ultracentrifugation is a method to determine the size distribution of macromolecules in a solute, cf. [7, 41]. The physical model leads to a Fredholm integral equation

$$f(y) = (\mathcal{T} g)(y) = \int g(x) K(x, y) dx, \tag{7.5}$$

whose kernel K of the integral operator $\mathcal{T}$ represents the sedimentation profile and is only implicitly given through the solution of the Lamm equation, a differential equation discussed in [34]. From the experimentally observed signal f, one must deduce the particles' or macromolecules' size distribution g. However, this is an ill-posed problem and requires regularization. State of the art regularizations for this problem are Tikhonov and maximum entropy regularization in [11, 32, 41]. Both methods have also been used in combination with Bayesian priors [7].

Partial information about a solute is often available. We consider the case in which we know a-priori that the solute is well separated into molecules of very different sizes. In other words, the seeked size distribution g is sparse, i.e., has only few peaks and is almost zero elsewhere. The sparser the expected distribution the smaller we may want to chose q. However, there is a trade off, because we then only minimize up to a constant factor. It seems reasonable to believe that heuristics can be developed to chose a near optimal q for a given experiment.

Before we apply the proposed iterative scheme to solve the analytical ultracentrifugation problem, we observe that the size distribution g must be nonnegative. We first discretize (7.5) by sampling on a finite grid. By using the nonnegativity as an additional regularization, we modify the application of the shrinkage process S_α in such a way that negative arguments are not shrinked in its original sense, but simply set to zero. It does not introduce any additional discontinuities, because $\varrho(x,\alpha) \to 0$ as $0 \leq x \to 0$. This procedure enforces a nonnegative limit.

In our numerical experiments, we consider different values of $0 < q < 1$ and compare the results in terms of how much sparsity we obtain while only introducing a relatively small residual $\|f - \mathcal{T}g\|$. We finally compare these findings to maximum entropy regularization that was used in [7, 41] to solve the ill-posed problem of analytical ultracentrifugation. We aim to verify that our proposed scheme can provide sparser solutions with sharper peaks, higher resolution and smaller residual.

7.3 Numerical Results

Maximum entropy regularization

$$\arg\min_g \left(\|f - \mathcal{T}g\|_{\ell_2}^2 + \beta \sum_n g_n \ln(g_n) \right)$$

is the state of the art tool to solve (7.5) for the analytical ultracentrifugation, cf. [7, 41]. It has been implemented in the softwaretool SEDFIT [41], that we use as a reference. SEDFIT uses f-statistics to choose β.

Sharper spikes, fewer nonzero entries, and higher resolution:
Our scheme is applied to a highly pure IgG antibody solute. Due to the purity, the "correct" solution to the underlying Fredholm integral equation must be highly sparse. We use 100 measurements on an equidistant grid to solve the discrete analogue of the integral equation (7.5). The residual $\|f - \mathcal{T}g_S\|_{\ell_2}$ of the SEDFIT solution g_S is 0.7878. Although the solution seems sparse, cf. the red graph in Figure 5, it has many small entries and the spikes are relatively wide.

The choices $0 < q < 1$ promote sparsity and, for sufficiently small $\alpha > 0$, our proposed scheme leads to smaller residuals, sharper spikes, and few small

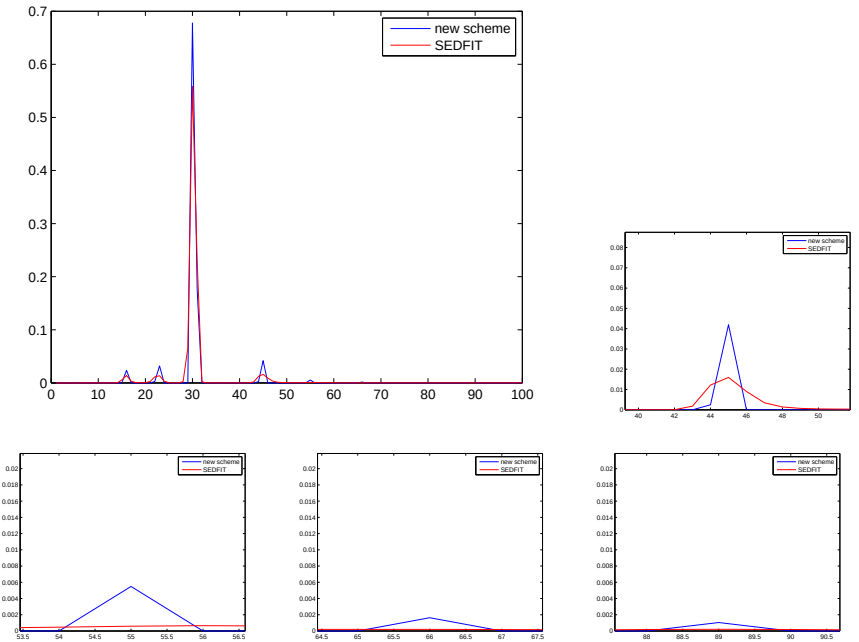


Figure 5. Comparison between SEDFIT and our proposed scheme with $q = 0.3$: we obtain sharper peaks, especially around 16 and 23. While SEDFIT is nonzero between 42 and 90, the solution to our scheme is zero except for peaks at 45, 55, 66, and 89.

entries, cf. Figure 5. The SEDFIT solution is nonzero between 42 and 90. The solution to our proposed scheme has peaks at 45, 55, 66, and 89, and vanishes in between. To verify that these peaks reflect the antibody solute (i.e., the peaks are real), we compute the maximum entropy solution for 1000 measurements on an equidistant grid, cf. Figure 6. This SEDFIT solution at this higher resolution has peaks around 45, 55, 66, and 89. Thanks to the sparsity promoting feature of our scheme, we "see" these peaks already with the much broader resolution of only 100 measurements. We observe that starting the iteration with the SEDFIT solution rather than the zero vector in (7.3) still leads to the same solution which indicates an intrinsic stability of the proposed scheme.

These results suggest that our proposed scheme has great potential when samples can be assumed to be highly pure. The method then leads to sharper spikes, fewer nonzero entries, and higher resolution.

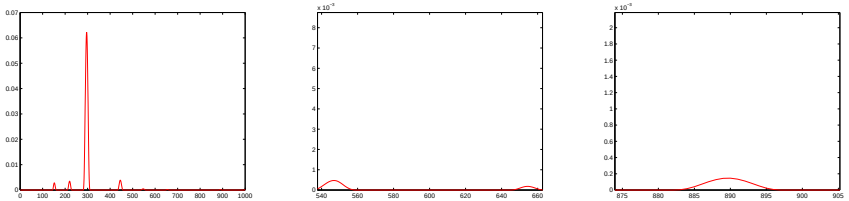


Figure 6. SEDFIT solution (maximum entropy regularized) for 1000 data points. There are peaks at 550, 660, and 890 (to compare with 100 data points, x-axes needs to be divided by 10 and y-axes needs to be multiplied by 10). Thus the peaks of our proposed method at 55, 66, and 89 are real and have the correct amplitude. They therefore reflect the antibody solute while using only a tenth of the data.

8 Conclusion

We have addressed variational problems with ℓ_q-constraints for $q \in (0,1)$. In case that computation time is crucial as it is in any real-time and on-line application, there are no sufficiently fast algorithms to solve them. By considering minimization up to a constant factor, we have overcome this limitation. We avoid costly iterative schemes and derive closed formulas for such minimizers. This approach provides a tool which makes problems for $q < 1$ more feasible than until now. If exact solutions are required, those minimizers can initialize iterative schemes to speed up their convergence and to find an accurate local minimum.

We have then modified the Landweber iteration with shrinkage applied at each iteration step in [14] by replacing the shrinkage rule with $\varrho_{h,s}^{(q)}(v_n, \alpha_n |v_n|^{q-1})$ to cover $q \in (0,1)$ as well. The proposed scheme has been used to solve the ill-posed problem of analytic ultracentrifugation. The results have been compared to the standard regularization for the analytical ultracentrifugation introduced in [7, 41]. We have verified that our proposed scheme can provide sparser solutions with sharper peaks, higher resolution and smaller residual. Thus, the scheme provides a useful add-on to the standard maximum entropy regularization.

It is known though that iterative schemes of the type presented in Section (7.1) converge relatively slowly and thus the computation time of our scheme is orders higher than those in [7, 41]. To present a competitive approach that can be incorporated into online applications such as the software-package SEDFIT/SEDPHAT [41], the method still needs a major tune up to derive a faster convergence, cf. [15] for possible directions.

For further theoretical foundation, it remains to find general conditions on L and on the bi-frame such that $\widetilde{F}^* L^\# L F$ is bounded on $\ell_q^{(\alpha_n)}$ and to compute the

difference between $\varrho_{h,s}^{(q)}$ and the exact minimizer of (4.1). It also remains to precisely determine the arising constants. For the analytical ultracentrifugation, the brute force discretization by means of sampling must still be replaced with a proper discretization scheme involving suitable ansatz functions and smoothness spaces. It also remains to verify that the Landweber iteration with q-dependent shrinkage converges towards a minimizer up to a certain error. We plan to address these topics in a forthcoming paper.

Acknowledgments. The author is grateful to Dr. Peter Schuck for providing the analytical ultracentrifugation data and for many insightful discussions.

Bibliography

[1] A. Antoniadis and J. Fan, Regularization of wavelet approximations, *J. Amer. Statist. Assoc.* **96** (2001), 939–967.

[2] R. Balan, J. Rosca and S. Rickard, Equivalence Principle for Optimization of Sparse versus Low-Spread Representations for Signal estimation in Noise, *Int. J. Imag. Syst. Tech.* **15** (2005), 10–17.

[3] T. Blumensath and M. E. Davies, Iterative hard thresholding for compressed sensing, *Appl. Comput. Harmon. Anal.* **27** (2009), 265–274.

[4] T. Blumensath, M. Yaghoobi and M. E. Davies, Iterative hard thresholding and L0 regularization, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing* **3** (2007), 877–880.

[5] T. Bonesky, K. Bredies, D. A. Lorenz and P. Maass, A generalized conditional gradient method for nonlinear operator equations with sparsity constraints, *Inverse Problems* **23** (2007), 2041–2058.

[6] T. Bonesky, S. Dahlke, P. Maass and T. Raasch, Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems, *Adv. Comput. Math., in press* (2010).

[7] P. H. Brown, A. Balbo and P. Schuck, A bayesian approach for quantifying trace amounts of antibody aggregates by sedimentation velocity analytical ultracentrifugation., *AAPS J.* **10** (2008), 481–493.

[8] E. Candès, J. Romberg and T. Tao, Stable signal recovery from incomplete and inaccurate measurements, *Comm. Pure Appl. Math.* **59** (2006), 1207–1223.

[9] A. Chambolle, R. A. DeVore, N. Y. Lee and B. J. Lucier, Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage, *IEEE Trans. Image Process.* **7** (1998), 319–335.

[10] R. Chartrand, Exact Reconstruction of Sparse Signals via Nonconvex Minimization, *IEEE Signal Processing Letters* **14** (2007), 707–710.

[11] D. J. Cox, Computer simulation of sedimentation in the ultracentrifuge. IV. Velocity sedimentation of self-associating solutes., *Arch. Biochem. Biophys.* **129** (1969), 106–123 (eng).

[12] S. Dahlke, M. Fornasier and T. Raasch, Adaptive Frame Methods for Elliptic Operator Equations, *Adv. Comput. Math.* **27** (2007), 27–63.

[13] S. Dahlke, T. Raasch, M. Werner, M. Fornasier and R. Stevenson, Adaptive frame methods for elliptic operator equations: the steepest descent approach, *IMA Journal of Numerical Analysis* **27** (2007), 717–740.

[14] I. Daubechies, M. Defrise and C. DeMol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Comm. Pure Appl. Math.* **57** (2004), 1413–1541.

[15] I. Daubechies, M. Fornasier and I. Loris, Accelerated Projected Gradient Method for Linear Inverse Problems with Sparsity Constraints, *J. Fourier Anal. Appl.* **14** (2008), 764–792.

[16] I. Daubechies and B. Han, Pairs of dual wavelet frames from any two refinable functions, *Constr. Approx.* **20** (2000), 325–352.

[17] D. Donoho, M. Elad and V. N. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, *IEEE Trans. Inform. Theory* **52** (2006), 6–18.

[18] D. Donoho and I. M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, *Biometrika* **81** (1994), 425–455.

[19] M. Ehler, On multivariate compactly supported bi-frames, *J. Fourier Anal. Appl.* **13** (2007), 511–532.

[20] M. Ehler, Nonlinear approximation associated with nonseparable wavelet bi-frames, *J. Approx. Theory* **161** (2009), 292–313.

[21] M. Ehler, The multiresolution structure of pairs of dual wavelet frames for a pair of Sobolev spaces, *Jaen J. Approx.* **2** (2010).

[22] M. Ehler and B. Han, Wavelet bi-frames with few generators from multivariate refinable functions, *Appl. Comput. Harmon. Anal.* **25** (2008), 407–414.

[23] M. Ehler and K. Koch, The Construction of Multiwavelet Bi-Frames and Applications to Variational Image Denoising, *Int. J. Wavelets, Multiresolut. Inf. Process.* **8** (2010), 431–455.

[24] M. Elad and M. Aharon, Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries, *IEEE Trans. Image Process.* **15** (2006), 3736–3745.

[25] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of inverse problems*, Mathematics and its applications v. 375, Kluwer Academic Publishers, Dordrecht, 1996.

[26] M. T. Figueiredo and R. D. Nowak, Wavelet-based image estimation: an empirical Bayes approach using Jeffrey's noninformative prior, *IEEE Trans. Image Process.* **10** (2001), 1322–1331.

[27] M. Fornasier, Domain decomposition methods for linear inverse problems with sparsity constraints, *Inverse Problems* **23** (2007), 2505–2526.

[28] M. Fornasier and H. Rauhut, Recovery Algorithms for Vector Valued Data with Joint Sparsity Constraints, *SIAM J. Numer. Anal.* **46** (2008), 577–613.

[29] H. Y. Gao, Wavelet shrinkage denoising using the non-negative garotte, *J. Comput. Graph. Statist.* **7** (1998), 469–488.

[30] H. Y. Gao and Andrew G. Bruce, Waveshrink with firm shrinkage, *Statistica Sinica* **7** (1997), 855–874.

[31] B. Han and Z. Shen, Dual Wavelet Frames and Riesz Bases in Sobolev Spaces, *Constr. Approx.* **29** (2009), 369–406.

[32] P. C. Hansen, Numerical Tools for the Analysis and solution of Fredholm integral-equations of the 1st kind, *Inverse Problems* **8** (1992), 849–872.

[33] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM, Philadelphia, 1998.

[34] O. Lamm, Die Differentialgleichung der Ultrazentrifugierung, *Ark. Mat. Astr. Fys.* **21B** (1929), 1–4.

[35] D. A. Lorenz, Convergence rates and source conditions for Tikhonov regularization with sparsity constraints, *J. Inv. Ill-Posed Problems* **16** (2008), 463–478.

[36] L. B. Montefusco and S. Papi, A parameter selection method for wavelet shrinkage denoising, *BIT* **43** (2003), 611–626.

[37] P. Mrázek, J. Weickert and G. Steidl, Correspondences between wavelet shrinkage and nonlinear diffusion, in: *Scale Space Methods in Computer Vision. Lecture Notes in Computer Science* (M. Lillholm L. D. Griffin, ed.), pp. 101–116, 2003.

[38] A. Niedermeier, E. Romaneessen and S. Lehner, Detection of coastlines in SAR images using wavelet methods, *IEEE Trans. Geosci. Remote Sensing* **38** (2000), 2270–2281.

[39] R. Ramlau and G. Teschke, A Tikhonov-based projection iteration for nonlinear ill-posed problems with sparsity constraints, *Numerische Mathematik* **104** (2006), 177–203.

[40] A. Ron and Z. Shen, Affine systems in $L_2(\mathbb{R}^d)$ II: dual systems, *J. Fourier Anal. Appl.* **3** (1997), 617–637.

[41] P. Schuck, Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling., *Biophys. J.* **78** (2000), 1606–1619.

[42] R. Stevenson and M. Werner, Computation of differential operators in aggregated wavelet frame coordinates, *IMA Journal of Numerical Analysis* **28** (2008), 354–381.

[43] T. Tao and B. Vidakovic, Almost everywhere convergence of general wavelet shrinkage estimators, *Appl. Comput. Harmon. Anal.* **9** (2000), 72–82.

[44] O. Yilmaz, R. Saab, R. Abugharbieh and M. McKeown, Underdetermined anechoic blind source separation via ℓ^q-basis-pursuit with $q < 1$, *IEEE Trans. Signal Process.* **55** (2007), 4004–4017.

[45] C. A. Zarzer, On Tikhonov regularization with non-convex sparsity constraints, *Inverse Problems* **25** (2009).

[46] Xiao-Ping Zhang and Mita D. Desai, Adaptive Denoising Based on SURE Risk, *IEEE Signal Processing Letters* **5** (1998), 265–267.

Author information

Martin Ehler, Section on Medical Biophysics, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, 9 Memorial Drive, Bethesda, MD 20892
and
Norbert Wiener Center, Department of Mathematics, University of Maryland, College Park, MD 20742, USA.
E-mail: ehlermar@mail.nih.gov, ehlermar@math.umd.edu